Decoherence, a Dynamical Approach to Superselection Rules ?

Domenico Giulini

Theoretische Physik, Universität Zürich,
Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Abstract. It is well known that the dynamical mechanism of decoherence may cause apparent superselection rules, like that of molecular chirality. These 'environment-induced' or 'soft' superselection rules may be contrasted with 'hard' superselection rules, like that of electric charge, whose existence is usually rigorously demonstrated by means of certain symmetry principles. We address the question of whether this distinction between 'hard' and 'soft' is well founded and argue that, despite first appearance, it might not be. For this we first review in detail some of the basic structural properties of the spaces of states and observables in order to establish a fairly precise notion of superselection rules. We then discuss two examples: 1.) the Bargmann superselection rule for overall mass in ordinary quantum mechanics, and 2.) the superselection rule for charge in quantum electrodynamics.

1 Introduction

To explain the (apparent) absence of interferences between macroscopically interpretable states – like states describing spatially localized objects – is the central task for any attempt to resolve the measurement problem. First attempts in this direction just imposed additional rules, like that of the Copenhagen school, who *defined* a measurement device as a system whose state-space is classical, in the sense that the superposition principle is fully broken: superpositions between any two states simply do not exist. In a more modern language this may be expressed by saying that any two states of such a system are *disjoint*, i.e., separated by a superselection rule (see below). Proper quantum mechanical systems, which in isolation do obey the superposition principle, can then inherit superselection rules when coupled to such classical measurement devices.

Whereas there can be no doubt that the notion of classicality, as we understand it here, is mathematically appropriately encoded in the notions of *disjointness* and *superselection rules*, there still remains the physical question how these structures come to be imposed. In particular, if one believes that fundamentally all matter is described by some quantum theory, there is no room for an independent classical world. Classicality should be a feature that is emerging in accordance with, and not in violation of, the basic rules of quantum mechanics. This is the initial credo of those who believe in the program of *decoherence* [16], which aims to explain classicality by means of

taking into account dynamical interactions with ambient systems, like the ubiquitous natural environment of the situation in question. (Note: It is not claimed to resolve the full measurement problem.) This leads to the notion of 'environment-induced superselection rules' [40].

Fundamental to the concept of dynamical decoherence is the notion of 'delocalization' [24]. The intuitive idea behind this is that through some dynamical process certain state characteristics ('phase relations'), which were locally accessible at one time, cease to be locally accessible in the course of the dynamical evolution. Hence locally certain superpositions cannot be verified anymore and an apparent obstruction to the superposition principle results. Such mechanisms are considered responsible for the above mentioned environment-induced superselection rule, of which a famous physical example is that of molecular chirality (see e.g. [38] and references therein). It has been established in many calculations of realistic situations that such dynamical processes of delocalization can be extremely effective over short time scales. But it is also intuitively clear that, mathematically speaking, it will never be strict in any finite time. Hence one will have to deal with notions of approximate- respectively asymptotic (for $t \to \infty$) superselection rules and disjointness of states [31,26], which needs some mathematical care.

Since for finite times such dynamical superselection rules are only approximately valid, they are sometimes called 'soft'. In contrast, 'hard' superselection rules are those which are rigorously established mathematical results within the kinematical framework of the theory, usually based on symmetry principles (see section 3 below), or on first principles of local QFT, like in the proof for the superselection rule for electric charge [33]. Such presentations seem to suggest that there is no room left for a dynamical origin of 'hard' superselection rules, and that hence these two notions of superselection rules are really distinct. However, we wish to argue that at least *some* of the existing proofs for 'hard' superselection rules give a false impression, and that quite to the contrary they actually need some dynamical input in order to be physically convincing. We will look at the case of Bargmann's superselection rule for total mass in ordinary quantum mechanics (which is clearly more of an academic example) and that of charge in QED. The discussion of the latter will be heuristic insofar as we will pretend that QED is nothing but quantum mechanics (in the Schrödinger representation) of the infinite-dimensional (constrained) Hamiltonian system given by classical electrodynamics. For a brief but general orientation on the subject of superselection rules and the relevant references we refer to Wightman's survey [39].

Let us stress again that crucial to the ideas presented here is of course that 'delocalized' does not at all mean 'destroyed', and that hence the loss of quantum coherence is only an *apparent* one. This distinction might be considered irrelevant FAPP (for all practical purposes) but it is important in attempts to understand apparent losses of quantum coherence *within* the standard dynamical framework of quantum mechanics.

As used here, the term 'local' usually refers to locality in the (classical) configuration space Q of the system, where we think of quantum states in the Schrödinger representation, i.e., as L^2-functions on Q. Every parametrization of Q then defines a partition into 'degrees of freedom'. Locality in Q is a more general concept than locality in ordinary physical space, although the latter forms a particular and physically important special case. Moreover, on a slightly more abstract level, one realizes that the most general description of why decoherence appears to occur is that only a *restricted* set of so-called physical observables are at ones disposal, and that *with respect to those* the relevant 'phase relations' *apparently* fade out of existence. It is sometimes convenient to express this by saying that decoherence occurs only with respect (or relative) to a 'choice' of observables [27]. Clearly this 'choice' is not meant to be completely free, since it has to be compatible with the dynamical laws and the physically realizable couplings (compare [24]). (In this respect the situation bears certain similarities to that of 'relevant' and 'irrelevant' degrees of freedom in statistical mechanics.) But to fully control those is a formidable task – to put it mildly. In any case it will be necessary to assume some a priori characterizations of what mathematical objects correspond to observables, and to do this in such a general fashion that one can effectively include superselection rules. This will be done in the next section.

2 Elementary Concepts

In this section we wish to convey a feeling for some of the concepts underlying the notion of superselection rules. We will take some care and time to do this, since many misconceptions can (and do!) arise from careless uses of these concepts. To gain intuition it is sometimes useful to dispense with some technicalities associated with infinite dimensions and continuous spectra and just look at finite dimensional situations; we will follow this strategy where indicated. We use the following, generally valid notations: $\mathcal{H}$ denotes a Hilbert space, $B(\mathcal{H})$ the algebra of bounded operators on $\mathcal{H}$. The antilinear operation of taking the hermitean conjugate is denoted by $*$ (rather than $\dagger$), which makes $B(\mathcal{H})$ a $*$-algebra. Given a set $\{A_\lambda\}$ where $\lambda \in \Lambda$ (= some index set), then by $\{A_\lambda\}'$ we denote the 'commutant' of $\{A_\lambda\}$ in $B(\mathcal{H})$, defined by

$$\{A_\lambda\}' := \{B \in B(\mathcal{H}) \,|\, BA_\lambda = A_\lambda B,\ \forall \lambda \in \Lambda\}. \tag{1}$$

Note that if the set $\{A_\lambda\}$ is left invariant under the $*$-map (in this case we call the set 'self-adjoint'), then $\{A_\lambda\}'$ is a $*$-subalgebra of $B(\mathcal{H})$. Also, the definition (1) immediately implies that

$$\mathcal{A} \subseteq \mathcal{B} \Rightarrow \mathcal{B}' \subseteq \mathcal{A}'. \tag{2}$$

2.1 Superselection Rules

There are many different ways to give a structural definition of superselection rules. Some stress the notion of *state* others the notion of *observable*. Often this dichotomy seems to result in very different attitudes towards the fundamental significance of superselection rules. This really seems artificial in a quantum mechanical context. In quantum *field* theory, i.e., if the underlying classical system has infinitely many degrees of freedom, the situation seems more asymmetric. This is partly due to the mathematical difficulties to define the full analog of the Schrödinger representation, i.e., to just construct the Hilbert space of states as L^2 space over the classical configuration space. In this paper we will partly ignore this mathematical difficulty and proceed heuristically by assuming that such a Schrödinger representation (of QED) exists to some level of rigour.

In traditional quantum mechanics, which stresses the notion of state, a system is fundamentally characterized by a Hilbert space, $\mathcal{H}$, the vectors of which represent (pure) states. We say 'represent' because this labeling by states through vectors is redundant: non-zero vectors which differ by an overall complex number label the same state, so that states can also be labeled by rays. We will use $\mathcal{P}\mathcal{H}$ to denote the space of rays in $\mathcal{H}$. In many cases of interest this Hilbert space is of course just identified with the space of L^2-functions over the classical configuration space. Now, following the original definition given by W^3 [35], we say that a superselection rule operates on $\mathcal{H}$, if not all rays represent pure states, but only those which lie entirely in certain mutually orthogonal subspaces $\mathcal{H}_i \subset \mathcal{H}$, where

$$\mathcal{H} = \bigoplus_i \mathcal{H}_i. \tag{3}$$

The only rays which correspond to pure states are those in the disjoint union

$$\bigcup_i \mathcal{P}\mathcal{H}_i\,. \tag{4}$$

Since no vector which lies skew to the partition (3) can, by assumption, represent a pure state, the superposition principle must be restricted to the $\mathcal{H}_i$. Moreover, since observables map pure states to pure states, they must leave the $\mathcal{H}_i$ invariant and hence all matrix-elements of observables between vectors from different sectors vanish. The $\mathcal{H}_i$ are called *coherent sectors* if the observables act irreducibly on them, i.e., if no further decomposition is possible; this is usually implied if a decomposition (3) is written down. States which lie in different coherent sectors are called *disjoint*. Note that disjointness of states is essentially also a statement about observables, since it means orthogonality of the original states *and* the respective states created from those with *all* observables. The existence of disjoint states is the characteristic feature of superselection rules.

From this we see that a partition (3) into coherent sectors implies that the set of physical observables is strictly smaller than the set of all self-adjoint (w.l.o.g. bounded) operators on $\mathcal{H}$. It can be characterized by saying that observables are those self adjoint operators on $\mathcal{H}$ which commute with the orthogonal projectors $P_i : \mathcal{H} \to \mathcal{H}_i$. So the P_i are themselves observables and generate the center (see below) of the algebra of observables.

This suggests a 'dual', more algebraic way to look at superselection rules, which starts with the algebra of observables $\mathcal{O}$. Then superselection rules are said to occur if the algebra of observables, $\mathcal{O}$, – which we think of as being given by bounded operators on some Hilbert space $\mathcal{H}$[1] – has a non-trivial center $\mathcal{O}^c$. Recall that

$$\mathcal{O}^c := \{A \in \mathcal{O} \,|\, AB = BA,\ \forall B \in \mathcal{O}\}. \tag{5}$$

Suppose $\mathcal{O}^c$ is generated by self-adjoint elements $\{C_\mu, \mu = 1, 2, ..\}$ which have simultaneous eigenspaces $\mathcal{H}_i$, then the $\mathcal{H}_i$'s are just the coherent sectors. Indeed, as already remarked, matrix elements of operators from $\mathcal{O}$ between states from different coherent sectors (i.e. differing in the eigenvalue of at least one C_μ) necessarily vanish. Thus if ϕ_1 and ϕ_2 are two non-zero vectors from $\mathcal{H}_i$ and $\mathcal{H}_j$ with $i \neq j$, their superposition $\phi := \phi_i + \phi_2$ defines a state whose density matrix $\rho := P_\phi$ (=orthogonal projector onto the ray generated by ϕ) satisfies

$$\mathrm{tr}(\rho A) = \mathrm{tr}((\lambda_1 \rho_1 + \lambda_2 \rho_2)A), \quad \forall A \in \mathcal{O}, \tag{6}$$

where $\lambda_{1,2} = \|\phi_{1,2}\|^2 / \|\phi\|^2$ and $\rho_{1,2} = P_{\phi_{1,2}}$. This means that ρ is a non-pure state of $\mathcal{O}$, since it can be written as a non-trivial convex combination of other density matrices. Hence we come back to the statements expressed by (3) and (4). Also note the following: in quantum mechanics the decomposition of a non-pure density matrix as a convex combination of pure density matrices – the so-called extremal decomposition – is generically not unique, thus pre-

[1] In Algebraic Quantum Mechanics one associates to each quantum system an abstract C^*-algebra, $\mathcal{C}$, which is thought of as being the mathematical object that fully characterizes the system in isolation, i.e. its intrinsic or 'ontic' properties. But this is not yet what we call the algebra of observables. This latter algebra is not uniquely determined by the former. It is obtained by studying faithful representations of $\mathcal{C}$ in some Hilbert space $\mathcal{H}$, such that $\mathcal{C}$ can be identified with some subalgebra of $B(\mathcal{H})$ (the bounded operators on $\mathcal{H}$). This is usually done by choosing a reference state (positive linear functional) on $\mathcal{C}$ and performing the GNS construction. Then $\mathcal{C}$ inherits a norm which is used to close $\mathcal{C}$ (as topological space) in $B(\mathcal{H})$. It is this resulting algebra which corresponds to our $\mathcal{O}$. Technically speaking it is a von Neumann algebra which properly contains an embedded copy of $\mathcal{C}$. The added observables (those in $\mathcal{O} - \mathcal{C}$) do not describe intrinsic but *contextual* properties. For example, it may happen that $\mathcal{O}$ has non-trivial center whereas $\mathcal{C}$ doesn't. In this case the superselection rules described by $\mathcal{O}$ are contextual. See [31] for a more extended discussion of this point.

venting the (ignorance-) interpretation as statistical "mixtures".[2] However, for the special density matrices of the form $\rho = |\phi\rangle\langle\phi|$, where $|\phi\rangle \in \mathcal{H}$, the extremal decomposition *is* unique and given by $\phi = \sum_i \lambda_i P_{\phi_i}$, where ϕ_i is the orthogonal projection of ϕ into $\mathcal{H}_i$, P_{ϕ_i} the orthogonal projector onto ϕ_i's ray, and $\lambda_i = \|\phi_i\|^2/\|\phi\|^2$. This is the relevance of superselection rules for the measurement problem: to produce *unique* extremal decompositions – and hence statistical 'mixtures' in the proper sense of the word – into an ensemble of pure states. There is a long list of papers dealing with the mathematical problem of how superselection sectors can arise dynamically; see e.g. [19,30,2,26] and the more general discussions in [27,31].

2.2 Dirac's Requirement

Dirac was the first who spelled out certain rules concerning the spaces of states and observables [6]. He defined the notion of *compatible* (i.e., simultaneously performable) *observations*, which mathematically are represented by a set of commuting observables, and the notion of a *complete set* of such observables, which is meant to say that there is precisely one state for each set of simultaneous "eigenvalues". Starting from the hypothesis that states are faithfully represented by rays, Dirac deduced that a complete set of such mutually compatible observables existed. But this only makes sense if all the observables in question have purely discrete spectra.

In the general case one has to proceed differently: We heuristically define Dirac's requirement as the statement, that *there exists at least one complete set of mutually compatible observables* and show how it can be rephrased mathematically so that it applies to all cases. In doing this we essentially follow Jauch's exposition [22]. To develop a feeling for what is involved, we will first describe some of the consequences of Dirac's requirement in the most simplest case: a finite dimensional Hilbert space. We will use this insight to rephrase it in such a way to stay generally valid in infinite dimensions.

Gaining intuition in finite dimensions. So let $\mathcal{H}$ be an n-dimensional complex Hilbert-space, then $B(\mathcal{H})$ is the algebra of complex $n \times n$ matrices. Physical observables are represented by hermitean matrices in $B(\mathcal{H})$, but we will explicitly *not* assume the converse, namely that *all* hermitean matrices correspond to physical observables. Rather we assume that the physical observables are somehow given to us by some set $\mathcal{S}$ of hermitean matrices. This set does not form an algebra, since taking products and complex linear combinations does not preserve hermiticity. But for mathematical reasons it would be convenient to have such an algebraic structure, and just work with the algebra $\mathcal{O}$ generated by this set, called the *algebra of observables*. [Note

[2] Hence the term 'mixture' for a non-pure state is misleading since we cannot tell the components and hence have no ensemble interpretation. For this reason we will say 'non-pure state' rather than 'mixture'.

the usual abuse of language, since only the hermitean elements in $\mathcal{O}$ are observables.] But for this replacement of $\mathcal{S}$ by $\mathcal{O}$ to be allowed $\mathcal{S}$ must have been a set of hermitean matrices which is uniquely determined by $\mathcal{O}$, for otherwise we can not reconstruct the set $\mathcal{S}$ from $\mathcal{O}$. To grant us this mathematical convenience we assume that $\mathcal{S}$ was already maximal, i.e. that $\mathcal{S}$ already contains all the hermitean matrices that it generates. But we stress that there seems to be no obvious reason why in a particular practical situation the set of physically realizable observables should be maximal in this sense.

We may choose a set $\{O_1, \ldots O_m\}$ of hermitean generators of $\mathcal{O}$. Then $\mathcal{O}$ may be thought of as the set of all complex polynomials in these (generally non-commuting) matrices. But note that we need not consider higher powers than $(n-1)$ of each O_i, since each complex $n \times n$ matrix O is a zero of its own characteristic polynomial p_O, i.e. satisfies $p_O(O) = 0$, by the theorem of Cayley-Hamilton. Since this polynomial is of order n, O^n can be re-expressed by a polynomial in O of order at most $(n-1)$. For example, the $*$-algebra generated by a single hermitean matrix O can be identified with the set of all polynomials of degree at most $(n-1)$ and whose multiplication law is as usual, followed by the procedure of reducing all powers n and higher of O via $p_O(O) = 0$.

Now let $\{A_1, \cdots, A_m\} =: \{A_i\}$ be a complete set of mutually commuting observables. It is not difficult to show that there exists an observable A and polynomials p_i, $i = 1, \cdots, m$ such that $A_i = p_i(A)$ (see [20] for a simple proof). This actually means that the algebra generated by $\{A_i\}$ is just the n-dimensional algebra of polynomials of degree at most $n-1$ in A (see below for justification), which we call $\mathcal{A}$. This algebra is abelian, which is equivalently expressed by saying that $\mathcal{A}$ is contained in its commutant (compare (1)):

$$\mathcal{A} \subseteq \mathcal{A}' \qquad \boxed{\text{`}\mathcal{A} \text{ is abelian'}} \tag{7}$$

Now comes the requirement of completeness. In terms of A it is easy to see that it is equivalent to the condition that A has a simple spectrum (i.e. the eigenvalues are pairwise distinct). This has the following consequence: Let B be an observable that commutes with A, then B is also a function of A, i.e., $p_B(B) = A$ for some polynomial p_B. The proof is simple: We simultaneously diagonalize A and B with eigenvalues α_a and β_a, $a = 1, \cdots, n$. We wish to find a polynomial of degree $n-1$ such that $p_B(\alpha_a) = \beta_a$. Writing $p_B(x) = a_{n-1}x^{n-1} + \cdots + a_0$, this leads to a system of n linear equations ($\alpha_a^b :=$ b^{th} power of α_a)

$$\sum_{b=0}^{n-1} \alpha_a^b a_b = \beta_a, \qquad \text{for } a = 1, \cdots, n \tag{8}$$

for the n unknowns $(a_0, \cdots, a_{n-1})$. Its determinant is of course just the Vandermonde determinant for the n tuple $(\alpha_1, \cdots, \alpha_n)$:

$$\det\{\alpha_a^b\} = \prod_{a<b} (\alpha_a - \alpha_b), \tag{9}$$

which is non-zero if and only if (=iff) A's spectrum is simple. This implies that every observable that commutes with $\mathcal{A}$ is already contained in $\mathcal{A}$. (It follows from this that the algebra generated by $\{A_i\}$ is equal to, and not just a subalgebra of, the algebra generated by $\{A\}$, as stated above.) Since a $*$-algebra is generated by its self-adjoint elements (observables), $\mathcal{A}$ cannot be properly enlarged as abelian $*$-algebra by adding more commuting generators. In other words, $\mathcal{A}$ is *maximal*. Since $\mathcal{A}'$ is a $*$-algebra, this can be equivalently expressed by

$$\mathcal{A}' \subseteq \mathcal{A} \qquad \boxed{\text{`}\mathcal{A} \text{ is maximal'}} \tag{10}$$

Equations (7) and (10) together are equivalent to Dirac's condition, which can now be stated in the following form, first given by Jauch [22]: the algebra of observables $\mathcal{O}$ contains a maximal abelian $*$-subalgebra $\mathcal{A} \subseteq \mathcal{O}$, i.e.,

$$\boxed{\text{Dirac's requirement, } 1^{\text{st}} \text{ version: } \exists\, \mathcal{A} \subseteq \mathcal{O} \text{ satisfying } \mathcal{A} = \mathcal{A}'} \tag{11}$$

This may seem as if Dirac's requirement could be expressed in purely algebraic terms. But this is deceptive, since the very notion of 'commutant' (compare (1)) makes reference to the Hilbert space $\mathcal{H}$ through $B(\mathcal{H})$. Without further qualification the term 'maximal' always means maximal *in* $B(\mathcal{H})$.[3]

This reference to $\mathcal{H}$ can be further clarified by yet another equivalent statement of Dirac's requirement. Since $\mathcal{A}$ consists of polynomials in the observable A, which has a simple spectrum, the following is true: there exists a vector $|g\rangle \in \mathcal{H}$, such that for *any* vector $\phi \in \mathcal{H}$ there exists a polynomial p_ϕ such that

$$p_\phi(A)|g\rangle = |\phi\rangle. \tag{12}$$

Such a vector $|g\rangle$ is called a generating or *cyclic vector* for $\mathcal{A}$ in $\mathcal{H}$. The proof is again very simple: let $\{\phi_1, \cdots, \phi_n\}$ be the pairwise distinct, non-zero eigenvectors of A (with any normalization); then choose

$$|g\rangle = \sum_{i=1}^{n} |\phi_i\rangle. \tag{13}$$

Equation (12) now defines again a system of n linear equations for the n coefficients $a_{n-1}, \cdots, a_0$ of the polynomial p_ϕ, whose determinant is again the Vandermonde determinant (9) for the n eigenvalues $\alpha_1, \cdots, \alpha_n$ of A. Conversely, if A had an eigenvalue, say α_1, with eigenspace $\mathcal{H}_1$ of two or higher dimensions, then such a cyclic $|g\rangle$ cannot exist. To see this, suppose it did, and let $|\phi_1^\perp\rangle \in \mathcal{H}_1$ be orthogonal to the projection of $|g\rangle$ into $\mathcal{H}_1$. Then $\langle\phi_1^\perp|p(A)g\rangle = 0$ for all polynomials p. Thus $|\phi_1^\perp\rangle$ is unreachable, contradicting our initial assumption. Hence a simple spectrum of A is equivalent to the existence of a cyclic vector.

[3] The condition for an abelian $\mathcal{A} \subseteq \mathcal{O}$ to be maximal in $\mathcal{O}$ would be $\mathcal{A} = \mathcal{A}' \cap \mathcal{O}$. Such abelian subalgebras *always* exist (use Zorn's Lemma to show this), in contrast to those $\mathcal{A} \subseteq \mathcal{O}$ which satisfy the stronger condition to be maximal in the ambient algebra $B(\mathcal{H})$, which need not exist for a given $\mathcal{O} \subset B(\mathcal{H})$.

The general case. In infinite dimensions we have to care a little more about the topology on the space of observables, since here there are many inequivalent ways to generalize the finite dimensional case. The natural choice is the so-called 'weak topology', which is characterized by declaring that a sequence $\{A_i\}$ of observables converges to the observable A if the sequence $\langle\phi|A_i|\psi\rangle$ of complex numbers converges to $\langle\phi|\mathcal{A}|\psi\rangle$ for all $|\phi\rangle, |\psi\rangle \in \mathcal{H}$. Hence one also requires that the algebra of observables is weakly closed (i.e., closed in the weak topology). Such a weakly closed $*$-subalgebra of $B(\mathcal{H})$ is called a W^*- or von-Neumann-algebra (we shall use the first name for brevity).

A crucial and extremely convenient point is, that the weak topology is fully encoded in the operation of taking the commutant (see (1)), in the following sense: Let $\{A_\lambda\}$ be any subset of $B(\mathcal{H})$, then $\{A_\lambda\}'$ is automatically weakly closed (see [22] p 716 for a simple proof) and hence a W^*-algebra. Moreover, the weak closure of a $*$-algebra $\mathcal{A} \subseteq B(\mathcal{H})$ is just given by $\mathcal{A}''$ (the commutant of the commutant). Hence we can characterize a W^*-algebra purely in terms of commutants: $\mathcal{A}$ is W^* iff $\mathcal{A} = \mathcal{A}''$.

This allows to easily generalize the notion of 'algebra generated by observables': Let $\{O_\lambda\}$ be a set of self-adjoint elements in $B(\mathcal{H})$, then $\mathcal{O} := \{O_\lambda\}''$ is called the (W^*-) algebra generated by this set. This definition is natural since $\{O_\lambda\}''$ is easily seen to be the smallest W^*-algebra containing $\{O_\lambda\}$, for if $\{O_\lambda\} \subseteq \mathcal{B} \subseteq \mathcal{O}$ for some W^*-algebra $\mathcal{B}$, then taking the commutant twice yields $\mathcal{B} = \mathcal{O}$.[4]

Now we see that Dirac's requirement in the form (11) directly translates to the general case if all algebras involved (i.e. $\mathcal{A}$ and $\mathcal{O}$) are understood as W^*-algebras. Now we also know what a 'complete set of (bounded) commuting observables' is, namely a set $\{A_\lambda\} \subseteq B(\mathcal{H})$ whose generated W^*-algebra $\mathcal{A} := \{A_\lambda\}''$ is maximal abelian: $\mathcal{A} = \mathcal{A}'$. This latter condition is again equivalent to the existence of a cyclic vector $|g\rangle \in \mathcal{H}$ for $\mathcal{A}$, where in infinite dimensions the definition of cyclic is that $\{\mathcal{A}|g\rangle\}$ is *dense* in (rather than equal to) $\mathcal{H}$. It is also still true that there is an observable A such that all $\mathcal{A}_\lambda$ are functions (in an appropriate sense, not just polynomials of course) of A [34]. But since A's spectrum may be (partially) continuous, there is no direct interpretation of a 'simple' spectrum as in finite dimensions. Rather, one now defines simplicity of the spectrum of A by the existence of a cyclic vector for $\mathcal{A} = \{A\}''$.

Now we come to our final reformulation of Dirac's condition. Namely, looking at (11), we may ask whether we could not reformulate the existence of such a maximal abelian $\mathcal{A}$ purely in terms of the algebra of observables $\mathcal{O}$ alone. This is indeed possible. We have $\mathcal{A} \subseteq \mathcal{O} \Rightarrow \mathcal{O}' \subseteq \mathcal{A}' = \mathcal{A} \subseteq \mathcal{O}$, hence $\mathcal{O}' \subseteq \mathcal{O}$. Since $\mathcal{O} = \mathcal{O}''$ the last condition is equivalent to saying that $\mathcal{O}'$ is abelian ($\mathcal{O}' \subseteq \mathcal{O}''$), or to saying that $\mathcal{O}'$ is the center $\mathcal{O}^c$ of $\mathcal{O}$, since by (1) and

[4] Note: for any $M \subseteq B(\mathcal{H})$ definition (1) immediately yields $M \subseteq M''$ and hence $M' \supseteq M'''$ (by (2)). But also $M' \subseteq M'''$ (by replacing $M \to M'$); therefore $M' = M'''$ for any $M \subseteq B(\mathcal{H})$.

(5) the center can be written as $\mathcal{O}^c = \mathcal{O} \cap \mathcal{O}'$. Now, conversely, it was shown in [23] that an abelian $\mathcal{O}'$ implies the existence of a maximal abelian $\mathcal{A} \subseteq \mathcal{O}$. Hence we have the following alternative formulation of Dirac's requirement, first spelled out, independently of (11), by Wightman [37], who called it the 'hypothesis of commutative superselection rules':

$$\boxed{\text{Dirac's requirement, } 2^{\text{nd}} \text{ version: } \mathcal{O}' \text{ is abelian}} \tag{14}$$

There are several interesting ways to interpret this condition. From its derivation we know that it is equivalent to the existence of a maximal abelian $\mathcal{A} \subseteq \mathcal{O}$. But we can in fact make an apparently stronger statement, which also relates to the earlier footnote 3, namely: (14) is equivalent to the condition, that *any* abelian $\mathcal{A} \subseteq \mathcal{O}$ that is maximal in $\mathcal{O}$, i.e. satisfies $\mathcal{A} = \mathcal{A}' \cap \mathcal{O}$, is also maximal in $B(\mathcal{H})$.[5]

2.3 Dirac's condition and gauge symmetries

Another way to understand (14) is via its limitations on *gauge-symmetries*. To see this, we mention that any W^*-algebra is generated by its unitary elements. Hence $\mathcal{O}'$ is generated by a set $\{U_\lambda\}$ of unitary operators. Each U_λ commutes with *all* observables and therefore generates a one-parameter group of gauge-transformations. Condition (14) is then equivalent to saying that the total gauge group, which is generated by all U_λ, is *abelian*. Note also that an abelian $\mathcal{O}'$ implies that the gauge-algebra, $\{U_\lambda\}'' = \mathcal{O}'$, is contained in the observables, $\mathcal{O}' \subseteq \mathcal{O}'' = \mathcal{O}$, so that $\mathcal{O}' = \mathcal{O}^c$. From this one can infer the following central statement:

$$\boxed{\begin{array}{l}\text{Dirac's requirement implies that gauge- and}\\ \text{sectorial structures are fully determined by the}\\ \text{center } \mathcal{O}^c \text{ of the algebra of observables } \mathcal{O}.\end{array}} \tag{15}$$

To see in what sense this is true we remark that for W^*-algebras we can simultaneously diagonalize all observables in $\mathcal{O}^c$. That means that we can write $\mathcal{H}$ in an essentially unique way as direct integral over the real line of Hilbert spaces $\mathcal{H}(\lambda)$ using some (Lebesgue-Stieltjes-) measure σ:

$$\mathcal{H} = \int_{\mathbb{R}}^{\oplus} d\sigma(\lambda)\, \mathcal{H}(\lambda). \tag{16}$$

Operators in $\mathcal{O}$ respect this decomposition in the sense that each $O \in \mathcal{O}$ acts on $\mathcal{H}$ componentwise via some bounded operator $O(\lambda)$ on $\mathcal{H}(\lambda)$. If $O \in \mathcal{O}^c$

[5] Proof: We need to show that ($\mathcal{O}'$ abelian) $\Leftrightarrow$ ($\mathcal{A} = \mathcal{A}' \cap \mathcal{O} \Rightarrow \mathcal{A} = \mathcal{A}'$). '$\Rightarrow$': $\mathcal{O}'$ abelian implies $\mathcal{O}' \subseteq \mathcal{O}'' = \mathcal{O}$ and $\mathcal{A} \subseteq \mathcal{O}$ implies $\mathcal{O}' \subseteq \mathcal{A}'$, so that $\mathcal{O}' \subseteq \mathcal{A}' \cap \mathcal{O}$. Hence $\mathcal{A} = \mathcal{A}' \cap \mathcal{O}$ implies $\mathcal{O}' \subseteq \mathcal{A}$, which implies $\mathcal{A}' \subseteq \mathcal{O}'' = \mathcal{O}$, and hence $\mathcal{A} = \mathcal{A}'$. '$\Leftarrow$': ($\mathcal{A} = \mathcal{A}' \cap \mathcal{O} \Rightarrow \mathcal{A} = \mathcal{A}'$) is equivalent to $\mathcal{A}' \subseteq \mathcal{O}$, which implies $\mathcal{O}' \subseteq \mathcal{A}'' = \mathcal{A}$ and hence that $\mathcal{O}'$ is abelian.

then each $O(\lambda)$ is a multiple $\phi(\lambda) \in \mathbb{C}$ of the unit operator. Moreover, the set of all $\{\mathcal{O}(\lambda)\}$ induced from $\mathcal{O}$ for each fixed λ acts irreducibly on $\mathcal{H}(\lambda)$.[6] Hence, provided that Dirac's requirement is satisfied, (16) is the generally valid version of (3). The notion of disjointness now acquires an intuitive meaning: two states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are separated by a superselection rule (are disjoint), iff their component-state-functions $\lambda \to |\psi_1(\lambda)\rangle$ and $\lambda \to |\psi_2(\lambda)\rangle$ have disjoint support on $\mathbb{R}$ (up to measure-zero sets). Note that by spectral decomposition the superselection observables can be decomposed into the projectors in $\mathcal{O}^c$, which for (16) are all given by multiplications with characteristic functions $\chi(\lambda)$ for σ-measurable sets in $\mathbb{R}$.

Non-abelian gauge groups We have seen that the fulfillment of Dirac's requirement allows to give a full structural characterisation for the spaces of (pure) states and observables. How general is this result? Does it exclude cases of physical interest? At first glance this seems indeed to be the case: just consider a situations with non-abelian gauge groups; for example, the quantum mechanical system of $n > 2$ identical spinless particles with n-particle Hilbert space $\mathcal{H} = L^2(\mathbb{R}^{3n})$ on which the permutation group $G = S_n$ of n objects acts in the obvious way by unitary operators $U(g)$. That these particles are identical means that observables must commute with each $U(g)$. Without further restrictions on observables one would thus define $\mathcal{O} := \{U(g),\, g \in G\}'$. Hence $\mathcal{O}'$ is the W^*-algebra generated by all $U(g)$, which is clearly non-abelian, thus violating (14). But does this generally imply that general particle statistics cannot be described in a quantum-mechanical setting which fulfills Dirac's requirement? The answer to this question is 'no'. Let us explain why.

If we decompose $\mathcal{H}$ according to the unitary, irreducible representations of G we obtain ([10][14])

$$\mathcal{H} = \bigoplus_{i=1}^{p(n)} \mathcal{H}_i \,, \tag{17}$$

where i labels the $p(n)$ inequivalent, unitary, irreducible representations D_i of G of dimension d_i. Each $\mathcal{H}_i$ has the structure $\mathcal{H}_i \cong \mathbb{C}^{d_i} \otimes \tilde{\mathcal{H}}_i$, where G acts irreducibly via D_i on $\mathbb{C}^{d_i}$ and trivially on $\tilde{\mathcal{H}}_i$ whereas $\mathcal{O}$ acts irreducibly via some $*$-representation π_i on $\tilde{\mathcal{H}}_i$ and trivially on $\mathbb{C}^{d_i}$. π_i and π_j are inequivalent if $i \neq j$. Hence we see that $\mathcal{H}_i$ furnishes an irreducible representation for $\mathcal{O}$, iff $d_i = 1$, i.e., for the Bose and Fermi sectors only. Pure states from these sectors are just the rays in the corresponding $\mathcal{H}_i$. In contrast, for $d_i > 1$, given a non-zero vector $|\phi\rangle \in \tilde{\mathcal{H}}_i$, all non-zero vectors in the d_i-dimensional subspace $\mathbb{C}^{d_i} \otimes |\phi\rangle \subset \mathcal{H}_i$ define the *same* pure state, i.e., the same expectation-value-functional on $\mathcal{O}$. Furthermore, a vector in $\mathcal{H}_i \cong \mathbb{C}^{d_i} \otimes \tilde{\mathcal{H}}_i$ which is not a pure

[6] It is this irreducibility statement which depends crucially on the fulfillment of Dirac's requirement. In general, the $O(\lambda)$'s will act irreducibly on $\mathcal{H}(\lambda)$ for each λ, iff $\mathcal{O}^c$ is maximal abelian *in* $\mathcal{O}'$, i.e., iff $\mathcal{O}^c = (\mathcal{O}^c)' \cap \mathcal{O}'$. But we already saw that (14) also implies $\mathcal{O}^c = \mathcal{O}'$ so that this is fulfilled.

tensor product defines a non-pure state, since the restriction of $O \in \mathcal{O}$ to $\mathcal{H}_i$ is of the form $\mathbf{1} \otimes \tilde{O}$, which means that a vector in $\mathcal{H}_i$ defines a state given by the reduced density matrix obtained by tracing over the left (i.e. $\mathbb{C}^{d_i}$) state space. From elementary quantum mechanics we know that the resulting state is pure, iff the vector in $\mathcal{H}_i$ was a pure tensor product (i.e. of rank one). Hence in those $\mathcal{H}_i$ where $d_i > 1$ not all vectors correspond to pure states, and those which do represent pure states in a redundant fashion by higher dimensional subspaces, sometimes called 'generalized rays' in the older literature on parastatistics [28].

However, the factors $\mathbb{C}^{d_i}$ are completely redundant as far as physical information is concerned, which is already fully encoded in the irreducible representations π_i of $\mathcal{O}$ on $\tilde{\mathcal{H}}_i$; no further physical information is contained in d_i-fold repetitions of π_i. Hence we can define a new, truncated Hilbert space

$$\tilde{\mathcal{H}} := \bigoplus_{i=1}^{p(n)} \tilde{\mathcal{H}}_i \,. \tag{18}$$

This procedure has also been called 'elimination of the generalized ray' in the older literature on parastatistics [18] – see also [14] for a more recent discussion of this point. Since every pure state in $\mathcal{H}$ is also contained in $\tilde{\mathcal{H}}$, just without repetition, these two sets are called 'phenomenological equivalent' in the literature on QFT (e.g. in chapter 6.1.C of [4]). The point is that pure states are now faithfully labelled by rays in the $\tilde{\mathcal{H}}_i$ and that $\mathcal{O}'$ – where the commutant is now taken in $B(\tilde{\mathcal{H}})$ rather than $B(\mathcal{H})$ – is generated by $\mathbf{1}$ and the $p(n)$ (commuting!) projectors into the $\tilde{\mathcal{H}}_i$'s. Hence Dirac's requirement is satisfied. But clearly the original gauge group has no action on $\tilde{\mathcal{H}}$ anymore, but there is also no physical reason why one should keep it.[7] It served to define $\mathcal{O}$, but then only its irreducible representations π_i are of interest. Only a residual action of the center of G still exists, but the gauge group generated by the projectors into the $\tilde{\mathcal{H}}_i$ consists in fact of the continuous group of $p(n)$ copies of $U(1)$, one global phase change for each sector. Its meaning is simply to induce the separation into the different sectors $(\tilde{\mathcal{H}}_i, \pi_i)$, and that in accordance with Dirac's requirement.

To sum up, we have seen that even if a theory is initially formulated via non-abelian gauge groups, we can give it a physically equivalent formulation that has at most a residual abelian gauge group left and hence obeys Dirac's requirement. Hence the 'obvious' counterexamples to Dirac's requirement turn out to be harmless. This is generally true in quantum mechanics,

[7] In [10] Dirac's requirement together with the requirement that the physical Hilbert space must carry an action of the gauge group has been used to "prove" the absence of parastatistics. In our opinion there seems to be no physical reason to accept the second requirement and hence the "proof"; compare [18] and [14].

but in quantum field theory there are genuine possibilities to violate Dirac's condition which we will ignore here.[8]

3 Superselection Rules via Symmetry Requirements

The requirement that a certain group must act on the set of all physical states is often the (kinematical) source of superselection rules. Here I wish to explain the structure of this argument.

Note first that in quantum mechanics we identify the states of a closed system with rays and not with vectors which represent them (in a redundant fashion). It is therefore not necessary to require that a symmetry group G acts on the Hilbert space $\mathcal{H}$, but rather it is sufficient that it acts on $\mathcal{PH}$, the space of rays, via so-called ray-representations. Mathematically this is a non-trivial relaxation since not every ray-representation of a symmetry group G (i.e. preserving the ray products) lifts to a unitary action of G on $\mathcal{H}$. What may go wrong is not that for a given $g \in G$ we cannot find a unitary (or anti-unitary) operator U_g on $\mathcal{H}$; that is assured by Wigner's theorem (see [3] for a proof). Rather, what may fail to be possible is that we can choose the U_g's in such a way that we have an *action*, i.e., that $\mathsf{U}_{g_1}\mathsf{U}_{g_2} = U_{g_1 g_2}$. As is well known, this is precisely what happens for the implementation of the Galilei group in ordinary quantum mechanics. Without the admission of ray representations we would not be able to say that ordinary quantum mechanics is Galilei invariant.

To be more precise, to have a ray-representation means that for each $g \in G$ there is a unitary[9] transformation U_g which, instead of the usual representation property, are only required to satisfy the weaker condition

$$\mathsf{U}_{g_1}\mathsf{U}_{g_2} = \exp(i\xi(g_1, g_2))\,\mathsf{U}_{g_1 g_2}, \tag{19}$$

for some function $\xi : G \times G \to \mathbb{R}$, called multiplier exponent, satisfying[10]

$$\xi(1, g) = \xi(g, 1) = 0, \tag{20}$$

$$\xi(g_1, g_2) - \xi(g_1, g_2 g_3) + \xi(g_1 g_2, g_3) - \xi(g_2, g_3) = 0. \tag{21}$$

The second of these conditions is a direct consequence of associativity: $\mathsf{U}_{g_1}(\mathsf{U}_{g_2}\mathsf{U}_{g_3}) = (\mathsf{U}_{g_1}\mathsf{U}_{g_2})\mathsf{U}_{g_3}$. Obviously these maps project to an action of

[8] An abelian $\mathcal{O}'$ implies that $\mathcal{O}$ is a von Neumann algebra of type I (see [7], chapter 8) whereas truly infinite systems in QFT are often described by type III algebras.

[9] For simplicity we ignore anti-unitary transformations. They cannot arise if, for example, G is connected.

[10] The following conditions might seem a little too strong, since it would be sufficient to require the equalities in (20) and (21) only mod 2π; this also applies to (22). But for our application in section 4 it is more convenient to work with strict equalities, which in fact implies no loss of generality; compare [32].

G on $\mathcal{PH}$. Any other lift of this action on $\mathcal{PH}$ onto $\mathcal{H}$ is given by a redefinition $\mathsf{U}_g \to \mathsf{U}'_g := \exp(i\gamma(g))\mathsf{U}_g$, for some function $\gamma : G \to \mathbb{R}$ with $\gamma(1) = 0$, resulting in new multiplier exponents

$$\xi'(g_1, g_2) = \xi(g_1, g_2) + \gamma(g_1) - \gamma(g_1 g_2) + \gamma(g_2), \tag{22}$$

which again satisfy (20) and (21). The ray representations U and U' are then said to be equivalent, since the projected actions on $\mathcal{PH}$ are the same. We shall also say that two multiplier exponents ξ, ξ' are equivalent if they satisfy (22) for some γ.

We shall now see how the existence of inequivalent multiplier exponents, together with the requirement that the group should act on the space of physical states, may clash with the superposition principle and thus give rise to superselection rules. For this we start from two Hilbert spaces $\mathcal{H}'$ and $\mathcal{H}''$ and actions of a symmetry group G on $\mathcal{PH}'$ and $\mathcal{PH}''$, i.e., ray representations U' and U'' on $\mathcal{H}'$ and $\mathcal{H}''$ up to equivalences (22). We consider $\mathcal{H} = \mathcal{H}' \oplus \mathcal{H}''$ and ask under what conditions does there exist an action of G on $\mathcal{PH}$ which restricts to the given actions on the subsets $\mathcal{PH}'$ and $\mathcal{PH}''$. Equivalently: when is $\mathsf{U} = \mathsf{U}' \oplus \mathsf{U}''$ a ray representation of G on $\mathcal{H}$ for some choice of ray-representations U' and U'' within their equivalence class? To answer this question, we consider

$$\begin{aligned} \mathsf{U}_{g_1}\mathsf{U}_{g_2} &= (\mathsf{U}'_{g_1} \oplus \mathsf{U}''_{g_1})(\mathsf{U}'_{g_2} \oplus \mathsf{U}''_{g_2}) \\ &= \exp(i\xi'(g_1, g_2))\mathsf{U}'_{g_1 g_2} \oplus \exp(\xi''(g_1, g_2))\mathsf{U}''_{g_1 g_2} \end{aligned} \tag{23}$$

and note that this can be written in the form (19), for some choice of ξ', ξ'' within their equivalence class, iff the phase factors can be made to coincide, that is, iff ξ' and ξ'' are equivalent. This shows that there exists a ray-representation on $\mathcal{H}$ which restricts to the given equivalence classes of given ray representations on $\mathcal{H}'$ and $\mathcal{H}'$, iff the multiplier exponents of the latter are equivalent. Hence, if the multiplier exponents ξ' and ξ'' are *not* equivalent, the action of G cannot be extended beyond the disjoint union $\mathcal{PH}' \cup \mathcal{PH}''$. Conversely, *if* we require that the space of physical states must support an action of G, then non-trivial superpositions of states in $\mathcal{H}'$ and $\mathcal{H}''$ must be excluded from the space of (pure) physical states.

This argument shows that if we insist of implementing G as symmetry group, superselection rules are sometimes unavoidable. A formal trick to avoid them would be not to require G, but a slightly larger group, $\bar{G}$, to act on the space of physical states. $\bar{G}$ is chosen to be the group whose elements we label by (θ, g), where $\theta \in \mathbb{R}$, and the multiplication law is

$$\bar{g}_1\bar{g}_2 = (\theta_1, g_1)(\theta_2, g_2) = (\theta_1 + \theta_2 + \xi(g_1, g_2), g_1 g_2). \tag{24}$$

It is easy to check that the elements of the form $(\theta, 1)$ lie in the center of $\bar{G}$ and form a normal subgroup $\cong \mathbb{R}$ which we call Z. Hence $\bar{G}/Z = G$ but G

need not be a subgroup of $\bar{G}$. $\bar{G}$ is a central $\mathbb{R}$ extension[11] of G (see e.g. [32]). Now a ray-representation U of G on $\mathcal{H}$ defines a proper representation U of $\bar{G}$ on $\mathcal{H}$ by setting

$$U_{(\theta,g)} := \exp(i\theta)\mathsf{U}_g. \tag{25}$$

Then $\bar{G}$ is properly represented on $\mathcal{H}'$ and $\mathcal{H}''$ and hence also on $\mathcal{H} = \mathcal{H}' \oplus \mathcal{H}''$. The above phenomenon is mirrored here by the fact that Z acts trivially on $\mathcal{PH}'$ and $\mathcal{PH}''$ but non-trivially on $\mathcal{PH}$, and the superselection structure comes about by requiring physical states to be fixed points of Z's action.

4 Bargmann's Superselection Rule

An often mentioned textbook example where a particular implementation of a symmetry group allegedly clashes with the superposition principle, such that a superselection rule results, is Galilei invariant quantum mechanics (e.g. [9]; see also Wightman's review [39]). We will discuss this example in detail for the general multi-particle case. (Textbook discussions usually restrict to one particle, which, due to Galilei invariance, must necessarily be free.) It will serve as a test case to illustrate the argument of the previous chapter and also to formulate our critique. Its physical significance is limited by the fact that the particular feature of the Galilei group that is responsible for the existence of the mass superselection rule ceases to exist if we replace the Galilei group by the Poincaré group (i.e. it is unstable under 'deformations'). But this is not important for our argument. [12] Let now G be the Galilei group, an element of which is parameterized by $(R, \boldsymbol{v}, \boldsymbol{a}, b)$, with R a rotation matrix in $SO(3)$, $\boldsymbol{v}$ the boost velocity, $\boldsymbol{a}$ the spatial translation, and b the time translation. Its laws of multiplication and inversion are respectively given by

$$\begin{aligned} g_1 g_2 &= (R_1, \boldsymbol{v}_1, \boldsymbol{a}_1, b_1)(R_2, \boldsymbol{v}_2, \boldsymbol{a}_2, b_2) \\ &= (R_1 R_2\,,\ \boldsymbol{v}_1 + R_1 \cdot \boldsymbol{v}_2\,,\ a_1 + R_1 \cdot \boldsymbol{a}_2 + \boldsymbol{v}_1 b_2\,,\ b_1 + b_2), \qquad (26) \\ g^{-1} &= (R, \boldsymbol{v}, \boldsymbol{a}, b)^{-1} = (R^{-1},\ -R^{-1} \cdot \boldsymbol{v}\,,\ -R^{-1} \cdot (\boldsymbol{a} - \boldsymbol{v}b)\,,\ -b). \qquad (27) \end{aligned}$$

We consider the Schrödinger equation for a system of n particles of positions $\boldsymbol{x}_i$, masses m_i, mutual distances $r_{ij} := \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$ which interact via a Galilei-invariant potential $V(\{r_{ij}\})$, so that the Hamilton operator becomes $\mathsf{H} = -\hbar^2 \sum_i \frac{\Delta_i}{2m_i} + V$. The Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^{3n}, d^3\boldsymbol{x}_1 \cdots d^3\boldsymbol{x}_n)$.

G acts on the space $\{\text{configurations}\} \times \{\text{times}\} \cong \mathbb{R}^{3n+1}$ as follows: Let $g = (R, \boldsymbol{v}, \boldsymbol{a}, b)$, then $g(\{\boldsymbol{x}_i\}, t) := (\{R \cdot \boldsymbol{x}_i + \boldsymbol{v}t + \boldsymbol{a}\}\,,\ t + b)$. Hence G has the obvious

[11] Had we defined the multiplier exponents mod 2π (compare footnote 10) then we would have obtained a $U(1)$ extension, which would suffice so far. But in the next section we will definitively need the $\mathbb{R}$ extension as symmetry group of the extended classical model discussed there.

[12] In General Relativity, where the total mass can be expressed as a surface integral at 'infinity', the issue of mass superselection comes up again; see e.g. [15] and [8].

left action on complex-valued functions on $\mathbb{R}^{3n+1}$: $(g,\psi) \to \psi \circ g^{-1}$. However, these transformations do *not* map solutions of the Schrödinger equations into solutions. But, as is well known, this can be achieved by introducing an $\mathbb{R}^{3n+1}$-dependent phase factor (see e.g. [13] for a general derivation). We set $M = \sum_i m_i$ for the total mass and $\boldsymbol{r}_c = \frac{1}{M}\sum_i m_i \boldsymbol{x}_i$ for the center-of-mass. Then the modified transformation, T_g, which maps solutions (i.e. curves in $\mathcal{H}$) to solutions, is given by

$$\mathsf{T}_g\psi(\{\boldsymbol{x}_i\},t) := \exp\left(\tfrac{i}{\hbar}M[\boldsymbol{v}\cdot(\boldsymbol{r}_c-\boldsymbol{a}) - \tfrac{1}{2}\boldsymbol{v}^2(t-b)]\right)\ \psi(g^{-1}(\{\boldsymbol{x}_i\},t)). \quad (28)$$

However, due to the modification, these transformations have lost the property to define an action of G, that is, we do *not* have $\mathsf{T}_{g_1}\circ\mathsf{T}_{g_2} = \mathsf{T}_{g_1 g_2}$. Rather, a straightforward calculation using (26) and (27) leads to

$$\mathsf{T}_{g_1}\circ\mathsf{T}_{g_2} = \exp(i\xi(g_1,g_2))\,\mathsf{T}_{g_1 g_2}, \quad (29)$$

with non-trivial multiplier exponent

$$\xi(g_1,g_2) = \tfrac{M}{\hbar}(\boldsymbol{v}_1\cdot R_1\cdot\boldsymbol{a}_2 + \tfrac{1}{2}\boldsymbol{v}_1^2 b_2). \quad (30)$$

Although each T_g is a mapping of *curves* in $\mathcal{H}$, it also defines a unitary transformation on $\mathcal{H}$ itself. This is so because the equations of motion define a bijection between solution curves and initial conditions at, say, $t=0$, which allows to translate the map T_g into a unitary map on $\mathcal{H}$, which we call U_g. It is given by

$$\mathsf{U}_g\psi(\{\boldsymbol{x}_i\}) = \exp\left(\tfrac{i}{\hbar}M[\boldsymbol{v}\cdot(\boldsymbol{r}_c-\boldsymbol{a}) + \tfrac{1}{2}\boldsymbol{v}^2 b]\right)\exp(\tfrac{i}{\hbar}\mathsf{H}b)\psi(\{R^{-1}(\boldsymbol{x}_i-\boldsymbol{a}+\boldsymbol{v}b)\}), \quad (31)$$

and furnishes a ray-representation whose multiplier exponents are given by (30). It is easy to see that the multiplier exponents are non-trivial, i.e., not removable by a redefinition (22). The quickest way to see this is as follows: suppose to the contrary that they were trivial and that hence (22) holds with $\xi' \equiv 0$. Trivially, this equation will continue to hold after restriction to any subgroup $G_0 \subset G$. We choose for G_0 the abelian subgroup generated by boosts and space translations, so that the combination $\gamma(g_1)-\gamma(g_1 g_2)+\gamma(g_2)$ becomes symmetric in $g_1, g_2 \in G_0$. But the exponent (30) stays obviously asymmetric after restriction to G_0. Hence no cancellation can take place, which contradicts our initial assumption.

The same trick immediately shows that the multiplier exponents are inequivalent for different total masses M. Hence, by the general argument given in the previous chapter, if $\mathcal{H}'$ and $\mathcal{H}''$ correspond to Hilbert spaces of states with different overall masses M' and M'', then the requirement that the Galilei group should act on the set of physical states excludes superpositions of states of different overall mass. This is Bargmann's superselection rule.

I criticize these arguments for the following reason: The dynamical framework that we consider here treats 'mass' as parameter(s) which serves to

specify the system. States for different overall masses are states of *different* dynamical systems, to which the superposition principle does not even potentially apply. In order to investigate a possible violation of the superposition principle, we must find a dynamical framework in which states of different overall mass are states of the *same* system; in other words, where mass is a dynamical variable. But if we enlarge our system to one where mass is dynamical, it is not at all obvious that the Galilei group will survive as symmetry group. We will now see that in fact it does not, at least for the simple dynamical extension which we now discuss.

The most simple extension of the classical model is to maintain the Hamiltonian, but now regarded as function on an extended, $6n + 2n$ - dimensional phase space with extra 'momenta' m_i and conjugate generalized 'positions' λ_i. Since the λ_i's do not appear in the Hamiltonian, the m_i's are constants of motion. Hence the equations of motion for the $\boldsymbol{x}_i$'s and their conjugate momenta $\boldsymbol{p}_i$ are unchanged (upon inserting the integration constants m_i) and those of the new positions λ_i are

$$\dot{\lambda}_i(t) = \frac{\partial V}{\partial m_i} - \frac{\boldsymbol{p}_i^2}{2m_i^2}, \tag{32}$$

which, upon inserting the solutions $\{\boldsymbol{x}_i(t), \boldsymbol{p}_i(t)\}$, are solved by quadrature.

Now, the point is that the new Hamiltonian equations of motion do not allow the Galilei group as symmetries anymore. But they do allow the $\mathbb{R}$-extension $\bar{G}$ as symmetries [13]. Its multiplication law is given by (24), with ξ as in (30). The action of $\bar{G}$ on the extended space of $\{\text{configurations}\}\times\{\text{times}\}$ is now given by

$$\begin{aligned}&\bar{g}(\{\boldsymbol{x}_i\}, \{\lambda_i\}, t) = (\theta, R, \boldsymbol{v}, \boldsymbol{a}, b)(\{\boldsymbol{x}_i\}, \{\lambda_i\}, t)\\ &= \left(\{R\boldsymbol{x}_i + \boldsymbol{v}t + \boldsymbol{a}\}\,, \{\lambda_i - (\tfrac{\hbar}{M}\theta + \boldsymbol{v}\cdot R\cdot \boldsymbol{x}_i + \tfrac{1}{2}\boldsymbol{v}^2 t)\}\,, t + b\right).\end{aligned} \tag{33}$$

With (24) and (30) it is easy to verify that this defines indeed an action. Hence it also defines an action on curves in the new Hilbert space $\bar{\mathcal{H}} := L^2(R^{4n}, d^{3n}\boldsymbol{x}d^n\lambda)$, given by

$$\bar{\mathsf{T}}_{\bar{g}}\psi := \psi \circ \bar{g}^{-1}\,, \tag{34}$$

which already maps solutions of the new Schrödinger equation to solutions, *without* invoking non-trivial phase factors. This is seen as follows: Let $\Psi(\{\boldsymbol{x}_i\}, \{\lambda_i\}, t) \in \bar{\mathcal{H}}$ and $\Phi(\{\boldsymbol{x}_i\}, \{m_i\}, t)$ its Fourier transform in the (λ_i, m_i) arguments:

$$\Phi(\{\boldsymbol{x}_i\}, \{\lambda_i\}, t) = (2\pi\hbar)^{-n/2}\int_{\mathbb{R}^n} d^n m\, \exp\left[\frac{i}{\hbar}\sum_{i=1}^{n} m_i\lambda_i\right]\, \Phi(\{\boldsymbol{x}_i\}, \{m_i\}, t). \tag{35}$$

For each set of masses $\{m_i\}$ the function $\Phi_{\{m_i\}}(\{\boldsymbol{x}_i\}, t) := \Phi(\{\boldsymbol{x}_i\}, \{m_i\}, t)$ satisfies the original Schrödinger equation. Since (34) does not mix different

sets of $\{m_i\}$ it induces a map $\bar{\mathsf{T}}_{\bar{g}}^{\{m_i\}}$ for each such set:

$$\begin{aligned}\bar{\mathsf{T}}_{\bar{g}}^{\{m_i\}}\Phi_{\{m_i\}}(\{\boldsymbol{x}_i\},t) := \exp\left[i\theta + \tfrac{i}{\hbar}M\left(\boldsymbol{v}\cdot(\boldsymbol{r}_c-\boldsymbol{a}) - \tfrac{1}{2}\boldsymbol{v}^2(t-b)\right)\right]\\ \times\, \Phi_{\{m_i\}}(g^{-1}(\{\boldsymbol{x}_i\},t))\end{aligned} \tag{36}$$

Via the Fourier transform (35) we represent $\bar{\mathcal{H}}$ as direct integral of $\mathcal{H}_{\{m_i\}}$'s, each of which isomorphic to our old $\mathcal{H} = L^2(\mathbb{R}^{3n}, d^3\boldsymbol{x}_1\cdots d^3\boldsymbol{x}_n)$, and on each of which (36) defines a unitary representation U of $\bar{G}$ the form (25) with U_g the ray-representation (31). This shows how the much simpler transformation law (34) contains the more complicated one (28) upon writing $\bar{\mathcal{H}}$ as a direct integral of vector spaces $\mathcal{H}_{\{m_i\}}$.

In the new framework the overall mass, M, is a dynamical variable, and it would make sense to state a superselection rule with respect to it. But now $\bar{G}$ rather than G is the dynamical symmetry group, which acts by a proper unitary representation on $\bar{\mathcal{H}}$, so that the requirement that the dynamical symmetry group should act on the space of physical states will now not lead to any superselection rule. Rather, the new and more physical interpretation of a possible superselection rule for M would be that we cannot localize the system in the coordinate conjugate to overall mass, which we call Λ, i.e., that only the *relative* new positions $\lambda_i - \lambda_j$ are observable.[13] (This is so because M generates translations of equal amount in all λ_i.) But this would now be a contingent physical property rather than a mathematical necessity. Note also that in our dynamical setup it is inconsistent to just state that M generates gauge symmetries, i.e. that Λ corresponds to a physically non existent degree of freedom. For example, a motion in real time along Λ requires a non-vanishing action (for non-vanishing M), due to the term $\int dt\, M\dot{\Lambda}$ in the expression for the action.

If decoherence were to explain the (ficticious) mass superselection rule, it would be due to a dynamical instability (as explained in [24]) of those states which are more or less localized in Λ. Mathematically this effect would be modelled by effectively removing the projectors onto Λ-subintervalls from the algebra of observables, thereby putting M (i.e. its projectors) into the center of $\mathcal{O}$. Such a non-trivial center should therefore be thought of as resulting from an approximation-dependent idealisation.

[13] A system $\{(\tilde{\lambda}_i, \tilde{m}_i\})$ of canonical coordinates including $M = \sum_i m_i$ is e.g. $\tilde{\lambda}_1 := \lambda_1$, $\tilde{m}_1 = M$ and $\tilde{\lambda}_i = \lambda_i - \lambda_1$, $\tilde{m}_i = m_i$ for $i = 2...n$. Then $\Lambda = \tilde{\lambda}_1$.

5 Charge Superselection Rule

In the previous case I said that superselection rules should be stated within a dynamical framework including as dynamical degree of freedom the direction generated by the superselected quantity. What is this degree of freedom in the case of a superselected electric charge and how does it naturally appear within the dynamical setup? What is its relation to the Coulomb field whose rôle in charge-decoherence has been suggested in [15]? In the following discussion I wish to investigate into these questions by looking at the Hamiltonian formulation of Maxwell's equation and the associated canonical quantization.

In Minkowski space, with preferred coordinates $\{x^\mu = (t, x, y, z)\}$ (laboratory rest frame), we consider the spatially finite region $Z = \{(t, x, y, z) : x^2 + y^2 + z^2 \leq R^2\}$. Σ denotes the intersection of Z with a slice $t = \text{const.}$ and $\partial\Sigma =: S_R$ its boundary (the laboratory walls). Suppose we wish to solve Maxwell's equations within Z, allowing for charged solutions. It is well known that in order for charged configurations to be stationary points of the action, the standard action functional has to be supplemented by certain surface terms (see e.g. [11]) which involve new fields on the boundary, which we call λ and f, and which represent a pair of canonically conjugate variables in the Hamiltonian sense. On the laboratory walls, $\partial\Sigma$, we put the boundary conditions that the normal component of the current and the tangential components of the magnetic field vanish. Then the appropriate boundary term for the action reads

$$\int_Z dt\, d\omega(\dot{\lambda} + \phi)f, \tag{37}$$

where ϕ is the scalar potential and $d\omega$ the measure on the spatial boundary 2-sphere rescaled to unit radius. Adding this to the standard action functional and expressing all fields on the spatial boundary by their multipole moments (so that integrals $\int_{\partial\Sigma} d\omega\, R^2$, $d\omega =$ measure on unit sphere, become $\sum_{lm}$), one arrives at a Hamiltonian function

$$H = \int_\Sigma \left[\tfrac{1}{2}(\boldsymbol{E}^2 + (\boldsymbol{\nabla} \times \boldsymbol{A})^2) + \phi(\rho - \boldsymbol{\nabla} \cdot \boldsymbol{E}) - \boldsymbol{A} \cdot \boldsymbol{j}\right] + \sum_{lm} \phi_{lm}(E_{lm} - f_{lm}). \tag{38}$$

Here the pairs of canonically conjugate variables are $(\boldsymbol{A}(\boldsymbol{x}), -\boldsymbol{E}(\boldsymbol{x}))$ and (λ_{lm}, f_{lm}), and E_{lm} are the multipole components of $\boldsymbol{n} \cdot \boldsymbol{E}$,

$$E_{lm} := \int_{\partial\Sigma} d\omega R^2\, Y_{lm} \boldsymbol{n} \cdot \boldsymbol{E}, \tag{39}$$

where $\boldsymbol{n}$ is the normal to $\partial\Sigma$. The scalar potential ϕ has to be considered as Lagrange multiplier. With the given boundary conditions the Hamiltonian is differentiable with respect to all the canonical variables[14] and leads to the

[14] This would not be true without the additional surface term (37). Without it one does not simply obtain the wrong Hamiltonian equations of motions, but

following equations of motion

$$\dot{\boldsymbol{A}} = \frac{\delta H}{\delta(-\boldsymbol{E})} = -\boldsymbol{E} - \boldsymbol{\nabla}\phi\,, \tag{40}$$

$$-\dot{\boldsymbol{E}} = -\frac{\delta H}{\delta \boldsymbol{A}} = \boldsymbol{j} - \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{A})\,, \tag{41}$$

$$\dot{\lambda}_{lm} = \frac{\partial H}{\partial f_{lm}} = -\phi_{lm}\,, \tag{42}$$

$$\dot{f}_{lm} = -\frac{\partial H}{\partial \lambda_{lm}} = 0\,. \tag{43}$$

These are supplemented by the equations which one obtains by varying with respect to the scalar potential ϕ, which, as already said, is considered as Lagrange multiplier. Varying first with respect to $\phi(\boldsymbol{x})$ (i.e. within Σ) and then with respect to ϕ_{lm} (i.e. on the boundary $\partial\Sigma$), one obtains

$$G(\boldsymbol{x}) := \boldsymbol{\nabla}\cdot\boldsymbol{E}(\boldsymbol{x}) - \rho(\boldsymbol{x}) = 0, \tag{44}$$

$$G_{lm} := E_{lm} - f_{lm} = 0. \tag{45}$$

These equations are constraints (containing no time derivatives) which, once imposed on initial conditions, continue to hold due to the equations of motion. [15]

This ends our discussion of the classical theory. The point was to show that it leaves no ambiguity as to what its dynamical degrees of freedom are, and that we had to include the variables λ_{lm} along with their conjugate momenta f_{lm} in order to gain consistency with the existence of charged configurations. The physical interpretation of the λ_{lm}'s is not obvious. Equation (42) merely relates their time derivative to the scalar potential's multipole moments on the boundary, which are clearly highly non-local quantities. The interpretation of the f_{lm}'s follow from (45) and the definition of E_{lm}, i.e. they are the multipole moments of the electric flux distribution $\varphi(\boldsymbol{n}) := R^2\boldsymbol{n}\cdot\boldsymbol{E}(R^2\boldsymbol{n})$. In particular, for $l = 0 = m$ we have

$$f_{00} = (4\pi)^{-\frac{1}{2}}\,Q, \tag{46}$$

none at all! Concerning the Langrangean formalism one should be aware that the Euler-Lagrange equations may formally admit solutions (e.g. with long-ranged (charged) fields) which are outside the class of functions which one used in the variational principle of the action (e.g. rapid fall-off). Such solutions are not stationary points of the action and their admittance is in conflict with the variational principle unless the expression for the action is modified by appropriate boundary terms.

[15] Equation (41) together with charge conservation, $\dot{\rho} + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$, shows that (44) is preserved in time, and (43) together with the boundary condition that $\boldsymbol{n}\cdot\boldsymbol{j}$ and $\boldsymbol{n}\times(\boldsymbol{\nabla}\times\boldsymbol{A})$ vanish on $\partial\Sigma$ show that (45) is preserved in time.

where Q is the total charge of the system. Hence we see that the total charge generates motions in λ_{00}. But this means that the degree of freedom labelled by λ_{00} truly exists (in the sense of the theory). For example, a motion along λ_{00} will cost a non-vanishing amount of action $\propto Q(\lambda_{00}^{\text{final}} - \lambda_{00}^{\text{initial}})$. A declaration that λ_{00} really labels only a gauge degree of freedom is *incompatible* with the inclusion of charged states. Similar considerations apply of course to the other values of l, m. But note that this conclusion is independent of the radius R of the spatial boundary 2-sphere $\partial\Sigma$. In particular, it continues to hold in the limit $R \to \infty$. We will not consistently get rid of physical degrees of freedom that way, even if we agree that realistic physical measurements will only detect field values in bounded regions of space-time. See [12] for more discussion on this point and the distinction between proper symmetries and gauge symmetries.

It should be obvious how these last remarks apply to the statement of a charge superselection rule. Without entering the technical issues (see e.g. [33]), its basic ingredient is Gauss' law (for operator-valued quantities), locality of the electric field and causality. That Q commutes with all (quasi-) local observables then follows simply from writing Q as surface integral of the local flux operator $R^2\boldsymbol{n} \cdot \hat{\boldsymbol{E}}$, and the observation that the surface may be taken to lie in the causal complement of any bounded space-time region. Causality then implies commutativity with any local observable.

In a heuristic Schrödinger picture formulation of QED one represents states Ψ by functions of the configuration variables $\boldsymbol{A}(\boldsymbol{x})$ and λ_{lm}. The momentum operators are obtained as usual:

$$-\boldsymbol{E}(\boldsymbol{x}) \longrightarrow -\mathrm{i}\frac{\delta}{\delta \boldsymbol{A}(\boldsymbol{x})}\,, \tag{47}$$

$$f_{lm} \longrightarrow -\mathrm{i}\frac{\partial}{\partial \lambda_{lm}}\,. \tag{48}$$

In particular, the constraint (45) implies the statement that on physical states Ψ we have[16]

$$\hat{Q}\Psi = -\mathrm{i}\sqrt{4\pi}\frac{\partial}{\partial \lambda_{00}}\Psi\,. \tag{49}$$

This shows that a charge superselection rule is equivalent to the statement that we cannot localize the system in its λ_{00} degree of freedom. Removing *by hand* the multiplication operator λ_{00} (i.e. the projectors onto λ_{00}-intervals) from our observables clearly makes Q a central element in the remaining algebra of observables. But what is the physical justification for this removal? Certainly, it is valid FAPP if one restricts to local observations in space-time. To state that this is a *fundamental* restriction, and not only an approximate

[16] Clearly all sorts of points are simply sketched over here. For example, charge quantization presumably means that λ_{00} should be taken with a compact range, which in turn will modify (48) and (49). But this is irrelevant to the point stressed here.

one, is equivalent to saying that for some fundamental reason we cannot have access to some of the *existing* degrees of freedom, which seems at odds with the dynamical setup. Rather, there should be a *dynamical* reason for why localizations in λ_{00} seem FAPP out of reach. The idea of decoherence would be that localizations in λ_{00} are highly unstable against dynamical decoherence.

We have mainly focussed on the charge superselection operator f_{00}, although the foregoing considerations make it clear that by the same argument any two different asymptotic flux distributions also define different superselection sectors of the theory. Do we expect these additional superselection rules to be physically real? First note that for $l > 0$ the f_{lm} are not directly related to the multipole moments of the charge distributions, as the latter fall-off faster than $\frac{1}{r^2}$ and are hence not detectable on the sphere at infinity. Conversely, the higher multipole moments f_{lm} are not measurable (in terms of electromagnetic fields) within any finite region of space-time, unlike the charge, which is tight to massive particles; any finite sphere enclosing all sources has the same total flux. But the f_{lm} can be related to the kinematical state of a particle through the retarded Coulomb field. In fact, given a particle with constant momentum $\boldsymbol{p}$, charge e and mass m, one obtains for the electric flux distribution at time t on a sphere centered at the instantaneous (i.e. at time t) particle position:[17]

$$\varphi_{\boldsymbol{p}}(\boldsymbol{n}) = \frac{em^2}{4\pi}\frac{[\boldsymbol{p}^2+m^2]^{\frac{1}{2}}}{[(\boldsymbol{p}\cdot\boldsymbol{n})^2+m^2]^{\frac{3}{2}}}. \tag{51}$$

Hence different incoming momenta would induce different flux distributions and therefore lie in different sectors. Given that these sectors exist this means

[17] Formula (51) requires a little more explanation: for a particle with general trajectory $\boldsymbol{z}(t)$ let t' be the retarded time for the space-time point $(\boldsymbol{x}, t)$, i.e., $t' = t - \|\boldsymbol{x} - z(t')\|$ ($c = 1$ in our units). Now we can use the well known formula for the retarded electric field (e.g. (14.14) in [21]) and compute the flux distribution on a sphere which lies in the space of constant time t, where it is centered at the retarded position $\boldsymbol{z}(t')$ of the particle. This flux distribution can be expressed as function of the retarded momentum $\boldsymbol{p}' := \boldsymbol{p}(t')$ and the retarded direction $\boldsymbol{n}' := [\boldsymbol{x} - \boldsymbol{z}(t')]/\|\boldsymbol{x} - \boldsymbol{z}(t')\|$ as follows ($E' := \sqrt{\boldsymbol{p}'^2 + m^2}$):

$$\varphi'_{\boldsymbol{p}'}(\boldsymbol{n}') = \frac{em^2}{4\pi}\frac{1}{[E' - \boldsymbol{p}'\cdot\boldsymbol{n}']^2}. \tag{50}$$

If the particle moves with *constant* velocity $\boldsymbol{v} := \dot{\boldsymbol{z}}$, the expression for the retarded Coulomb field can be rewritten in terms of the instantaneous position $\boldsymbol{z}(t)$ by using $\boldsymbol{z}(t) = \boldsymbol{z}(t') + \boldsymbol{v}\|\boldsymbol{x} - \boldsymbol{z}(t')\|$. With respect to this center it is purely radial. Then one calculates the flux distribution on a sphere which again lies in the space of constant time t, but now centered at $\boldsymbol{z}(t)$ rather than $\boldsymbol{z}(t')$. This function can be expressed in terms of the instantaneous direction $\boldsymbol{n} := [\boldsymbol{x} - \boldsymbol{z}(t)]/\|\boldsymbol{x} - \boldsymbol{z}(t)\|$ and the instantaneous momentum $\boldsymbol{p} := \boldsymbol{p}(t)$. One obtains (51).

that different incoming momenta cannot be coherently superposed and no incoming localized states be formed, unless one also adds the appropriate incoming infrared photons to just cancel the difference of the asymptotic flux distributions. This is achieved by imposing the 'infrared coherence condition of Zwanziger [41][18] the effect of which is to 'dress' the charged particles with infrared photons which just subtract their retarded Coulomb fields at large spatial distances. Hence coherent superpositions of particles with different momenta can only be formed if they are dressed by the right amount of incoming infrared photons.

As a technical aside we remark that this can be done without violating the Gupta-Bleuler transversality condition $k^\mu a_\mu(k)|\text{in}\rangle = 0$ in the zero-frequency limit, precisely because of the surface term (37)[11]. This resolved an old issue concerning the compatibility of the infrared coherence condition on one hand, and the Gupta-Bleuler transversality condition on the other [17,42]. From what we said earlier concerning the consistency of the variational principle in the presence of charged states, such an apparent clash of these two conditions had to be expected: without the surface variables one cannot maintain gauge invariance at spatial infinity (i.e. in the infrared limit) and at the same time include charged states. In the charged sectors the longitudinal infrared photons correspond to real physical degrees of freedom and it will naturally lead to inconsistencies if one tries to eliminate them by imposing the Gupta-Bleuler transversality condition also in the infrared limit. However, a gauge symmetry in the infrared limit can be maintained if one adds the asymptotic degrees of freedom in the form of surface terms.

These remarks illustrate how the rich superselection structure associated with different asymptotic flux distributions f_{lm} renders the problem of characterizing state spaces in QED for charged sectors fairly complicated. This problem has been studied within various formalisms including algebraic QFT [5] and lattice approximations, where the algebra of observables can be explicitly presented [25]. However, all this takes for granted the existence of the superselection rules, whereas we would like to see whether they really arise from some physical impossibility to localize the system in the degrees of freedom labelled by λ_{lm}. What physics should prevent us from forming incoming localized wave packets of charged *undressed* (in the sense above) particles, which would produce coherent superpositions of asymptotic flux distributions from the sectors with $l \geq 1$? This cries out for a decoherence mechanism to provide a satisfying physical explanation. The case of charge superselection is, however, more elusive, since localizations in λ_{00} do not have an obvious physical interpretation. Compare the controversy between [1,29] on one side and [36] on the other.

[18] Basically it says that the incoming scattering states should be eigenstates to the photon annihilation operators $a^{\text{in}}_\mu(k)$ in the zero-frequency limit.

References

1. Aharonov, Y., Susskind, L (1967): Charge Superselection Rule. Phys. Rev. **155**, 1428-1431
2. Araki, H. (1980): A Remark on the Machida-Namiki Theory of Measurement. Prog. Theo. Phys. **64**, 719-730
3. Bargmann, V. (1964): Note on Wigner's Theorem on Symmetry Operations. Jour. Math. Phys. **5**, 862-868
4. Bogolubov, N.N., Logunov, A.A., Oksak, A.I., Todorov, I.T. (1990): *General Principles of Quantum Field Theory*, (Kluwer, Dordrecht).
5. Buchholz, D. (1982): The Physical State Space of Quantum Electrodynamics. Commun. Math. Phys. **85**, 49-71
6. Dirac, P.A.M. (1930): *The Principles of Quantum Mechanics* (Clarendon Press, Oxford)
7. Dixmier, J. (1981): *Von Neumann Algebras* (North Holland, Amsterdam)
8. Dominguez. A.E., Kozameh, C.N., Ludvigsen, M. (1997): Superselection Sectors in Asymptotic Quantization of Gravity, gr-qc/9609071
9. Galindo, A., Pascual, P. (1990): *Quantum Mechanics I* (Springer, Berlin)
10. Galindo, A., Morales, A., Nuñez-Lagos, R. (1962): Superselection Principle and Pure States of n-Identical Particles. Jour. Math. Phys. **3**, 324-328
11. Gervais, J.L., Zwanziger, D. (1980): Derivation From First Principles of the Infrared Structure of Quantum Electrodynamics. Phys. Lett. B **94**, 389-393
12. Giulini, D. (1995): Asymptotic Symmetry Groups of Long-Ranged Gauge Configurations. Mod. Phys. Lett A **10**, 2059–2070
13. Giulini, D. (1996): On Galilei Invariance and the Bargmann Superselection Rule. *Ann. Phys. (NY)* **249**, 222-235
14. Giulini, D. (1995): Quantum Mechanics on Spaces with Finite Fundamental Group. Helv. Phys. Acta **68**, 438–469
15. Giulini, D., Kiefer, C. Zeh, H.D. (1995): Symmetries, Superselection Rules, and Decoherence. Phys. Lett. A **199**, 291–298
16. Giulini, D, Joos, E., Kiefer, C., Kupsch, J., Stamatescu, I.-O., Zeh, H.-D. (1996): *Decoherence and the Appearance of a Classical World in Quantum Theory*, (Springer, Berlin)
17. Haller, K. (1978): Gupta-Bleuler condition and infrared-coherent states. Phys.Rev. D **18**, 3045–3051
18. Hartle, J.B., Taylor, J.R. (1969): Quantum Mechanics of Paraparticles. Phys. Rev. **178**, 2043-2051
19. Hepp, K. (1972): Quantum Theory of Measurement and Macroscopic Observables. Helv. Phys. Acta **45**, 237-248
20. Isham, C.J. (1995): *Lectures on Quantum Theory.* (Imperial College Press)
21. Jackson, J.D. (1975) *Classical Electrodynamics*, second edition, (John Wiley & Sons, New York)
22. Jauch, J.M. (1960): Systems of Observables in Quantum Mechanics. Helv. Phys. Acta **33**, 711–726
23. Jauch, J.M., Misra, B. (1961): Supersymmetries and Essential Observables. Helv. Phys. Acta **34**, 699–709
24. Joos, E. (2000): Elements of Environmental Decoherence. In:*Decoherence: Theoretical, Experimental and Conceptual Problems*, Lecture Notes in Physics Vol. 538 (Springer, Berlin), eds. P. Blanchard et al.

25. Kijowski, J., Rudolph, G., Thielmann, A. (1997): Algebra of Observables and Charge Superselection Sectors for QED on the Lattice. Commun. Math. Phys. **188**, 535-564
26. Kupsch, J. (2000): Mathematical Aspects of Decoherence. In: *Decoherence: Theoretical, Experimental and Conceptual Problems*, Lecture Notes in Physics Vol. 538(Springer, Berlin), eds. P. Blanchard et al.
27. Landsman, N.P. (1995): Observation and Superselection in Quantum Mechanics. Stud. Hist. Phil. Mod. Phys. **26**, 45-73
28. Messiah, A.M., Greenberg,O.W. (1964): Symmetrization Postulate and Its Experimental Foundation. Phys. Rev. **236**, B 248-B 267
29. Mirman, R (1969): Coherent Superposition of Charge States. Phys. Rev. **186**, 1380–1383
30. Pfeifer, P. (1980): A simple model for irreversible dynamics from unitary time evolution. Helv. Phys. Acta **53**, 410-415
31. Primas, H. (2000): Asymptotically disjoint quantum states. In: *Decoherence: Theoretical, Experimental, and Conceptual Problems*, Lecture Notes in Physics Vol. 538 (Springer, Berlin), eds. P. Blanchard et al.
32. Raghunathan, M.S. (1994): Universal Central Extensions. Rev. Math. Phys. **6**, 207–225
33. Strocchi, F, Wightman, A.S. (1974): Proof of the Charge Superselection Rule in Local Relativistic Quantum Field Theory. Jour. Math. Phys. **15**, 2198-2224; Erratum Ibid **17** (1976), 1930-1931
34. von Neumann, J. (1931): Über Funktionen von Funktionaloperatoren. Ann. Math. (Princeton) **32** 191-226
35. Wick, G.C., Wightman, A.S., Wigner, E.P. (1952): The Intrinsic Parity of Elementary Particles. Phys. Rev. **88**, 101–105
36. Wick, G.C., Wightman, A.S., Wigner, E.P. (1970): Superselection Rule for Charge. Phys. Rev. D **1** 3267–3269
37. Wightman, A.S. (1959): Relativistic Invariance and Quantum Mechanics (Notes by A. Barut). Nouvo Cimento, Suppl. **14**, 81-94
38. Wightman, A.S., Glance N. (1989): Superselection Rules in Molecules. Nucl. Phys. B (Proc. Suppl.) **6**, 202-206
39. Wightman, A.S. (1995): Superselection Rules; Old and New. Il Nuovo Cimento **110 B**, 751–769
40. Zuerk, W. (1982): Environment-Induced Superselection Rules. Phys. Rev. D **26**, 1862-1880
41. Zwanziger, D. (1976): Physical States in Quantum Electrodynamics. Phys. Rev. D **14**, 2570-2589
42. Zwanziger, D. (1978): Gupta-Bleuler and Infrared-Coherence Subsidiary Conditions. Phys. Rev. D **18**, 3051-3057